Agenda

- Challenges, Metrics, and Measurement Strategies
- Measurement Technologies
  - NetFlow
  - NBAR
  - SA Agent
  - ART MIB
- Case Studies

Traffic Engineering in a Multiservice Network

Network Must Provide Each Application With Different Service Level Characteristics Simultaneously
# Application Traffic Engineering Challenges

- **Planning**
  - Verify application impact on the network and vice-versa
  - Deployment of new applications and services
  - Gain understanding of traffic flows
  - Utilization and latency dependent on application mix

- **Troubleshooting**
  - Bursts in traffic load caused by applications
  - Service degradation
  - Rogue applications that hijack the network
  - Misconfigured applications

## Application Performance Metrics

- **Responsiveness**
- **Availability**
- **Throughput**
- **Utilization**
- **Jitter**
Responsiveness

- The response time to fulfill a request for service
- Measurement based on the start of fulfillment
  When the server responds with requested data
  Not all services requested are the same effort

Availability

- As a percentage of time the service is actually provided
  Network service is not busy
  Network service is not down
  Network service does not introduce errors
- As a function of MTBF and MTTR
Throughput

- Transmission rate of observed data transport measured in:
  - Bits/bytes per second
  - Packets per second
- Differences between “provisioned” and “observed” due to:
  - Oversubscription of circuit
  - Router utilization
  - Server congestion

Utilization

- Percent of total bandwidth consumed
  - Per interface
  - Per interface by traffic class
  - Per interface by application or protocol
Delay Variation—“Jitter”

- The change in packet interval for sequential packets
  - Source transmits with constant inter-packet interval
  - During network transit, inter-packet interval is affected
  - At destination, variation of inter-packet interval is jitter

Performance Measurement Strategies

- **Synthetic**
  - Sampling Method: Observed

- **Embeded Agents**
  - Collection Method: External Probes

- **Device/Link**
  - Scope of Measurement: End-to-End/Path

- **User**
  - Perspective of Measurement: Network
### Sampling Method

#### Synthetic
- **Definition**
  Actual end-user network traffic where performance is measured by timing specific application traffic flows.
- **Advantages**
  Most accurate for live application traffic on a specified link.
- **Disadvantages**
  Limited to measuring:
  - Existing traffic types, which may not be present on the network at all times.
  - Existing traffic patterns, which may not reflect patterns for new or future applications.

#### Observed
- **Definition**
  Network traffic generated strictly for the purpose of measuring a network performance characteristic.
- **Advantages**
  Measures performance:
  - Between any two points in the network.
  - Controllable, on a continuous basis.
  - By traffic class based on IP Precedence marking.
- **Disadvantages**
  Only an approximation for performance of live traffic.

### Collection Method

#### Embedded
- **Definition**
  Mechanisms for collection of network statistics are integrated into the network communication device (e.g., router or switch), itself.
- **Advantages**
  Follows network infrastructure.
  Gathers metrics that can not be observed externally.
- **Disadvantages**
  Performance monitoring has device-level performance implications.

#### External
- **Definition**
  Mechanisms for collection of network statistics are provided by a stand-alone device specifically designed to collect network performance statistics.
- **Advantages**
  Validation of performance performed independent of the devices that transmit network traffic.
- **Disadvantages**
  More hardware to administer.
  Observed statistics limited to points of deployment.
### Scope of Measurement

#### Device or Link Oriented
- **Definition**: Performance measurement based on analysis of specific device or device interface, and typically based on utilization rates.

#### End-to-End
- **Definition**: Performance measurement based on analysis of response time across two or more network devices, and typically based on latency.

#### Advantages
- **Device or Link Oriented**: Detailed application performance monitoring of critical network links.
- **End-to-End**: Starting point performance troubleshooting, reflects end-user experience.

#### Disadvantages
- **Device or Link Oriented**: When network-wide performance problems exist, how does one select which device or link to evaluate?
- **End-to-End**: Prior knowledge of relevant end-to-end paths is needed.

### Perspective of Measurement

#### User
- **Definition**: Measurement based on performance statistics measured at the end-user workstation.

#### Network
- **Definition**: Measurement based on performance statistics measured in network devices.

#### Advantages
- **User**: Accurate measurement of end-user experience.
- **Network**: Easy to deploy, and non-intrusive to the desktop.

#### Disadvantages
- **User**: Scale and distribution issues, intrusive on the desktop.
- **Network**: Imperfect understanding of end-user experience.
Application Measurement Technologies

Measurement Technologies

<table>
<thead>
<tr>
<th>NetFlow</th>
<th>ART MIB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEASURES:</strong> Device interface traffic rate by s/d IP address, port number or AS</td>
<td><strong>Application Response Time SNMP MIB</strong></td>
</tr>
<tr>
<td><strong>Sampling:</strong> Observed</td>
<td><strong>MEASURES:</strong> Response time of live application traffic to server device</td>
</tr>
<tr>
<td><strong>Collection:</strong> Embedded</td>
<td><strong>Sampling:</strong> Observed</td>
</tr>
<tr>
<td><strong>Scope:</strong> Device/Link</td>
<td><strong>Collection:</strong> External Probe</td>
</tr>
<tr>
<td><strong>Perspective:</strong> Network</td>
<td><strong>Scope:</strong> End-to-End</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SA Agent</th>
<th>NBAR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Service Assurance Agent</strong></td>
<td><strong>Network-Based Application Recognition</strong></td>
</tr>
<tr>
<td><strong>MEASURES:</strong> Latency and Jitter between source router and specified target</td>
<td><strong>MEASURES:</strong> Device interface utilization by application type or QoS traffic class</td>
</tr>
<tr>
<td><strong>Sampling:</strong> Synthetic</td>
<td><strong>Sampling:</strong> Observed</td>
</tr>
<tr>
<td><strong>Collection:</strong> Embedded</td>
<td><strong>Collection:</strong> Embedded</td>
</tr>
<tr>
<td><strong>Scope:</strong> End-to-End</td>
<td><strong>Scope:</strong> Device/Link</td>
</tr>
<tr>
<td><strong>Perspective:</strong> User/Network</td>
<td><strong>Perspective:</strong> Network</td>
</tr>
</tbody>
</table>
NetFlow Defined

- Flows are defined by 7 keys:
  - Source Address
  - Destination Address
  - Source Port
  - Destination Port
  - Layer 3 Protocol
  - TOS byte (DSCP)
  - Input Interface
- Flows are unidirectional
- Flows are enabled on a per input-interface basis
- Flows can be configured “on-demand” or continuous

Flow Data Exported to Management Application
Measurement Technologies
SA Agent

Service Assurance (SA) Agent Defined

- Embedded technology in Cisco IOS® devices
- Synthetic traffic for various protocols
- Supports IP precedence for QoS
- Measures latency, jitter and availability
## Service Assurance Agent Operation Types

<table>
<thead>
<tr>
<th>Operations</th>
<th>IOS Release 11.2/11.3</th>
<th>IOS Release 12.0</th>
<th>IOS Release 12.0T/12.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICMP Echo (Ping)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IP Path Echo</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>LU0/LU2 (3270 Ping)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SSCP (SNA)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>UDP Echo</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>TCP Connect</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>UDP Jitter</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>IP Precedence</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DNS Name Resolution</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DHCP Get IP Address</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DLSw Peer Tunnel</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>HTTP</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Note: IP precedence can be combined with other operation types to simulate QoS traffic marking.
NBAR
Network-Based Application Recognition

- Cisco IOS-based protocol discovery featuring:
  - Measure interface utilization by protocol or traffic class
  - “Stateful Inspection” for non-symmetric protocols
  - See list of protocols
  - Define QoS traffic classes
  - HTTP traffic by URL or MIME type

SNMP or Web Reporting
## NBAR

### Partial List of Protocols Identified

- **Statefully-inspected protocols…**
  - SQL*Net
  - Sun RPC
  - H.323
  - MSEXchange
  - UNIX r commands
  - VDOLive
  - RealAudio
  - MS Netshow
  - HTTP (via URL/MIME)
  - FTP
- **Static-port protocols…**
  - CuSeeMe
  - IPSec
  - SMTP
  - EGP
  - IRC
  - SNMP
  - GRE
  - L2TP
  - SOCKS
  - ICMP
  - LDAP
  - SSH
  - IGMP
  - NFS
  - Streamworks
  - IMAP
  - NNTP
  - TFTP
  - OSPF
  - NTP
  - Telnet
  - BGP
  - NetBIOS
  - X Windows
  - DHCP
  - Notes
  - Etc.
  - DNS
  - OSPF
  - Finger
  - PCAnywhere
  - Gopher
  - POP3
  - HTTP
  - RIP
  - RSVP
  - Secure HTTP

### Support for Additional Protocols Is Software Definable

- **Statefully-inspected protocols…**
  - SQL*Net
  - Sun RPC
  - H.323
  - MSEXchange
  - UNIX r commands
  - VDOLive
  - RealAudio
  - MS Netshow
  - HTTP (via URL/MIME)
  - FTP
- **L4–L7 protocols which dynamically assign TCP/UDP ports**
  - SQL*Net
  - Sun RPC
  - H.323
  - MSEXchange
  - UNIX r commands
  - VDOLive
  - RealAudio
  - MS Netshow
  - HTTP (via URL/MIME)
  - FTP
- **Protocol Discovery**
  - Graph
  - Table
  - View All
  - By output packet count
  - Refresh
  - Last refresh on 06/01/03

## NBAR

### Example of Protocol Discovery
Measurement Technologies
ART MIB

ART MIB Implementation Example

- Dedicated RMON probes for critical links and high-speed backbones
- ART monitoring option installed on key probes
- Application monitoring tools for measuring application performance and response time
ART MIB Functionality

- TCP protocols only (1.0)
- Based upon well-known destination port
- Default protocols:

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Default Protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOL</td>
<td>NNTP</td>
</tr>
<tr>
<td>COMPUSERV</td>
<td>NOTESTCP</td>
</tr>
<tr>
<td>DLSW RD</td>
<td>ORACLSQL</td>
</tr>
<tr>
<td>DLSW WR</td>
<td>REALAUD</td>
</tr>
<tr>
<td>DNS_TCP</td>
<td>SMTP</td>
</tr>
<tr>
<td>DOOM</td>
<td>SNA_TCP</td>
</tr>
<tr>
<td>FTP-CTRL</td>
<td>SOCKET</td>
</tr>
<tr>
<td>FTP-DATA</td>
<td>SQLNET_N</td>
</tr>
<tr>
<td>HTTP</td>
<td>SUNRPC_T</td>
</tr>
<tr>
<td>HTTPS</td>
<td>TELNET</td>
</tr>
<tr>
<td>NB_DGM_T</td>
<td>XWINDOW</td>
</tr>
<tr>
<td>NB_NS_T</td>
<td></td>
</tr>
<tr>
<td>NB_SSN_T</td>
<td></td>
</tr>
<tr>
<td>NEWS_TCP</td>
<td></td>
</tr>
</tbody>
</table>

Application Level Response Time

Client Latency            Server Latency

Network Flight Time

Example of Reporting

- Server address
- Number of clients
- For all (summary) clients:
  - Min/avg/max response time
  - Request/response pair count
  - Response time distribution
  - Client and server octets and overflows
  - Retries
  - Timeouts

Identify Application

Packet Level Measurement

Example: FTP

Response Time

SEQ 101
ACK 101
SEQ 102
SEQ 103
SEQ 104
ACK 104
SEQ 105
ACK 105
Case Study 1:

Troubleshooting Application Response Time Issues

Network Environment

- Large banking institution
  Web-based application being used for both internal and external customers
  Application had been running for several months without problems

- Organization
  Four service groups:
  - Firewall group
  - Local application group (proxy server)
  - Network group
  - Application group (application deployment)

- The network
  Clients attached to 10 Mbps VLANS
  Web server access via the Internet
  Web access via a proxy server
  Firewalls running NAT
**Problem Description**

- Upon switching to secured access (HTTPS), users experienced excessive response time problems (“hourglass mode” for minutes)

- All service groups were “pointing fingers” at each other

---

**Technical Analysis**

- Analyze network traffic
  - TrafficDirector and external probes with ART monitoring were used to profile the network
  - No errors were found on the network
  - Traffic volumes were within acceptable range
  - Client conversations could not be mapped due to NAT

- Analyze application flow
  - ART Monitor was used to determine response time at each step
  - Average response time
    - Probe 1: 130 ms
    - Probe 2: 135 ms
    - Probe 3: 138 ms
  - Response time seemed reasonable, yet the client was still experiencing problems
The Resolution

Problem Isolation
- Initial analysis showed no problems
- Probe 1 was moved between the client and proxy server
- Significant delays in response time between the client and proxy server
- TrafficDirector was used to capture and analyze traffic data

Resolution
- Reconfigured proxy server
- Response time for the user now acceptable

Summary
- ART Monitor provided insight into the user experience
- TrafficDirector provided the tools to analyze the details of the problem
- The groups were able to share information and resolve the problem

Case Study 2:
Analyzing Application-Network Interaction and QoS Activation
Network Environment

- Workstation configuration tracking
- E-mail/web browsing
- ERP

Problem Description

- Existing:
  PC configuration tracking
  Maintain consistent software configurations on 3,000+ PCs
  ERP
  Manage inventory worth over $400M at over 380 retail locations
  Manage daily pricing strategies at over 350 retail locations
  E-mail and web access
- Future:
  Voice over IP
  Use data infrastructure to reduce PSTN and cell phone charges

Problems:

PC configuration tracking congests T-1 Frame Relay from HQ to SP
ERP, e-mail and web traffic are deprived reasonable bandwidth
VoIP deployment to retail outlets is not possible
Home users dial-in on 800 number, then use expensive cell phones
Technical Analysis

- NBAR is used for:
  - Baseline application utilization
  - QoS traffic class definition
  - Monitoring QoS traffic class utilization

Solution: QoS Applied to Traffic Classes

<table>
<thead>
<tr>
<th>Traffic Class</th>
<th>Minimum Guarantee</th>
<th>Maximum Guarantee</th>
<th>Queue Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC Configuration</td>
<td>10% of Link</td>
<td>33% of Link</td>
<td>Normal</td>
</tr>
<tr>
<td>ERP</td>
<td>40% of Link</td>
<td>None</td>
<td>Normal</td>
</tr>
<tr>
<td>E-Mail and Web</td>
<td>10% of Link</td>
<td>20% of Link</td>
<td>Normal</td>
</tr>
<tr>
<td>VOIP</td>
<td>15Kbps</td>
<td>15Kbps</td>
<td>High</td>
</tr>
</tbody>
</table>
Application Traffic Engineering
Summary

• **Effective traffic engineering requires:**
  
  Determining which application performance metrics are relevant to the network environment
  
  Understanding which measurement strategies are most appropriate for obtaining performance metrics
  
  Knowledge of performance measurement technologies that will facilitate each measurement strategy

Other Networkers’ Topics of Interest...

• SLAM presentation
• SAA presentation
• Demo at booth
• CiscoWorks2000 Product Overview
Application Aware Traffic Engineering and Monitoring

Session 2603

Please Complete Your Evaluation Form

Session 2603